LIFAR, 7 janvier 1998

# Automates infinis

## Une approche combinatoire et analytique

## via les marches aléatoires

Cyril.Banderier@inria.fr

http://algo.inria.fr/banderier/

**Projet Algorithmes**
**INRIA-Rocquencourt (France)**

# Mots de Dyck et nombres de Catalan

Les mots de Dyck,
*i.e.* (les (mots (bien)) () parenthésés), sont facilement mis en bijection avec une marche sur $\mathbb{N}$ à pas $\pm 1$.
Combien y a-t-il de mots de longueur $n$ ?
Notons $c_n$ ce nombre et on forme la série génératrice

$$C(z) = \sum_{n \geq 0} c_n z^n$$

On peut directement lire la série génératrice sur la grammaire

$$S \to \nearrow S \searrow S \,|\, \epsilon \;\Rightarrow\; C(z) = z^2 C^2(z) + 1 \;\Rightarrow$$

$$C(z) = \frac{1 + \sqrt{1 - 4z^2}}{2z^2} = 1 + z^2 + 2z^4 + 5z^6 + 14z^8 + \dots$$

On obtient les nb de Catalan : $c_{2n} = \frac{\binom{2n}{n}}{n+1}$ $\square$

But : généralisation avec plus de pas possibles, en autorisant les multiplicités (*i.e.* on perd le déterminisme).

# Abondance de points (de vue) ne nuit pas

Les mots engendrés par une grammaire infinie "régulière" (alphabet infini, ensemble infini de productions)

$\approx$

traces d'une marche dans un automate infini

$\approx$

puissances itérées d'une matrice infinie (mais row finite)

$\approx$

parcours dans un arbre de possibilités

$\approx$

marches (aléatoires) sur un réseau discret.

Exemple avec les mots de Dyck :

- Grammaire infinie : $S_0 \to (1)S_1 \,|\, \epsilon$ et pour $i > 0$ : $S_i \to (i+1)S_{i+1} \,|\, (i-1)S_{i-1}$
- Marche sur $\mathbb{N}$ : $[(0), (k) \to (k-1)(k+1)]$
- $P$ matrice d'un graphe orientée $p_{i,j} = 1 \Rightarrow$ il y a un arc de $(i)$ à $(j)$. chemins de longueur $n$ : $M = P^n = (m_{i,j})_{i,j=0\ldots\infty}$ $\square$

## Wanted : $érie génératrice bivariée

Pour des règles fixées, on va expliciter les séries génératrices des "marches aléatoires sur les entiers" qui décomptent les chemins de longueur $n$ (les mots de longueurs $n$) ou les excursions (les mots qui commencent et finissent par la même lettre).

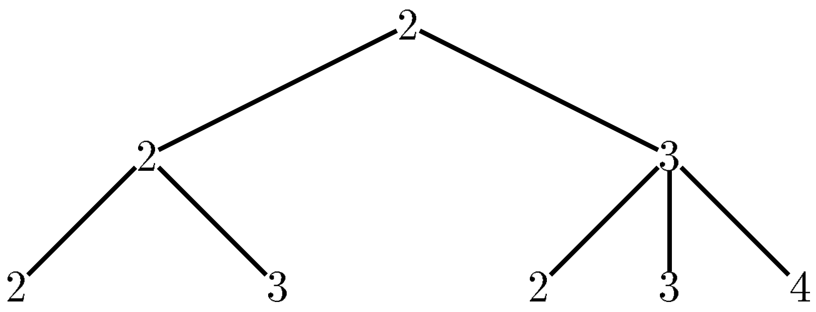
$$F(z,u) = \sum_{n\geq 0, k\geq 0} f_{n,k} u^k z^n$$

•$f_{n,k}$ est le nombre de chemins de longueur $n$ qui vont de l'état initial à l'état $k$.
•$F(z,1)$ compte tous les chemins.
•$F(z,0)$ compte les excursions.
•$[u^k]F(z,u)$ compte les chemins de longueur $n$, finissant en $k$.

# Les *ECO*systèmes

confer le papier (nov.98) de C. Banderier, M. Bousquet-Mélou, A. Denise, P. Flajolet, D. Gardy et D. Gouyou-Beauchamps :
"On generating functions on generating trees".

Arbre de possibilités :
un nœud étiqueté $k$ a $k$ fils.




Conjecture de Pinzani :
règle de récriture $(k) \rightarrow e_1(k) \ldots e_k(k)$
"Poids linéaire" $\Rightarrow$ s.g. rationnelle ?
"Pas avant bornés" $\Rightarrow$ s.g. algébrique ?

# Critique de la rationalité linéaire

Définition. Le poids d'une règle de récriture $[(k), \rightarrow e_1(k) \ldots e_k(k)]$ est défini par $p(k) := e_1(k) + e_2(k) + \ldots e_k(k)$.

**Théorème 1.** *Une règle de poids linéaire a une série génératrice rationnelle.*

Preuve [Denise98]. Posons $p(k) = \alpha k + \beta$. Soient $l_1, l_2, \ldots l_{f_n}$ les étiquettes des $f_n$ nœuds de profondeur $n$, on a

$$f_{n+2} = (\alpha l_1 + \beta) + (\alpha l_2 + \beta) + \ldots + (\alpha l_{f_n} + \beta)$$

$$= \alpha f_{n+1} + \beta f_n$$

d'où

$$F(z) = \frac{z + (f_0 - \alpha)z^2}{1 - \alpha z - \beta z^2}.$$

$\square$

*n.b. :* On peut obtenir une "généralisation" de ce résultat avec le point de vue opérateur exposé plus loin.

# Primalité et rationalité

**Exemple 1.** *La marche* $(k) \to (k+1)(k)(k+3)(2)^{k-3}$ *a une s.g. rationnelle rationnelle.*

Preuve. Simple application du théorème précédent

$$F(z) = \frac{z + (f_0 - \alpha)z^2}{1 - \alpha z - \beta z^2}.$$

$\square$

Les nombres premiers sont loin d'être un langage rationnel ou même algébrique (densité...) donc il est contre-intuitif que la modification suivante de la marche précédente puisse donner une s.g. rationnelle :

**Application 1.** *La marche*

$$\begin{cases} (k) \to (k+1)(k)(k+3)(2)^{k-3} & \textit{si } k \textit{ non premier} \\ (k) \to (k+1)(k+1)(k+2)(2)^{k-3} & \textit{si } k \textit{ premier} \end{cases}$$

*a une série génératrice rationnelle.*

Preuve. Dans ce cas aussi le poids vaut $p(k) = 5k - 2$ donc $F(z,1) = \frac{z-3z^2}{1-5z+2z^2}$. $\square$

# Marches dans un borné

**Théorème 2.** *La série génératrice d'une marche dans un borné de $\mathbb{Z}^n$ est rationnelle.*

Preuve. Transformation en automate puis théorème de Chomsky-Schützenberger □

On a une asymptotique exacte :

$$[z^n]F(z) = \sum_{j=1}^{m} P_j(n)\rho_j^{-n}$$

# Marches à pas bornés

**Théorème 3.** *La série génératrice d'une marche (dans $\mathbb{Z}$) à pas bornés est rationnelle.*

Preuve. $f_n(u) = L^n f_0(u)$,
donc si $L = P(u)$ est un polynôme, on a :

$$F(z,u) = \frac{1}{1 - P(u)z}.$$

*n.b.* : $u = 0 \Rightarrow$ résidus $\Rightarrow$ s.g. algébrique. □

# La méthode nucléaire

La méthode du noyau fait partie du folklore des 70's (proba, combinatoire [*e.g.* Knuth I.2.2.1]).

**Problème.** *On veut déterminer $F$, on a $F.N = Q$ où $Q$ dépend de $n$ inconnues.*

Solution : on résout $N = 0$, ce qui donne $n$ racines, qu'on réinjecte pour obtenir un système $n \times n$. Et alors $F = N/Q$. $\square$

**Application.** A partir du système infini

$$\begin{cases} f_0(u) = 1 \\ \dots \\ f_{n+1}(u) = M f_n(u) \end{cases}$$

en multipliant par $z^{n+1}$ et comme $M = L - \{u^{<0}\}L$ est linéaire, en sommant on a :

$$\sum_{n \geq 0} z^{n+1} f_{n+1}(u) = 1 + z \sum_{n \geq 0} z^n M f_n(u)$$

$$F(z,u) = 1 + zLF(z,u) - z\{u^{<0}\}LF$$

Et on applique la méthode du noyau à :

$$F(z,u)(1 - P(u)) = 1 - z \sum_{i=0}^{c-1} p_i(u) \partial_i F(z,0)$$

# Marches sur $\mathbb{N}$ à pas bornés

$f_{n+1} = Lf_n(u) - \{u^{<0}\}Lf_n(u)$

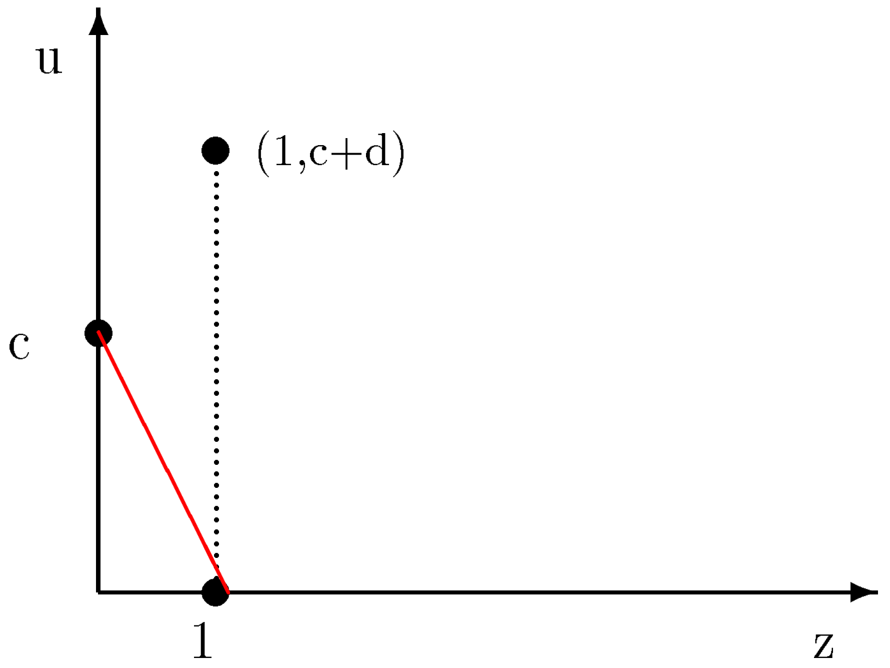
$$F(z,u)(1 - zP(u)) = 1 - z\sum_{j=0}^{c-1} c_j(u)\partial_u^j F(z,0)$$

Posons

$$Q(u,z) := u^c\left(1 - z\sum_{j=1}^{c-1} c_j(u)\partial_u^j F(z,0)\right).$$

et le noyau est

$$N(u,z) := u^c - zu^cP(u)$$



polygone de Newton pour développer les solutions de $N = 0$ en série de Puiseux, on obtient $c + d$ solutions $u_1, \ldots u_{c+d}$, dont :

- $c$ petites branches que nous noterons $u_1, \ldots, u_c$ ;
- $d$ branches infinies en 0, que nous noterons $v_1, \ldots, v_d$

En regardant $N$ et $Q$ comme des polynômes en $u$ et en remplaçant $u$ par $u_i$, on voit que les $u_i$ sont aussi racines de $Q(u, z)$ (car $FN = Q$).

**Théorème 4.** *La s.g. d'une marche sur $\mathbb{N}$ à pas bornés est algébrique.*

Preuve.

$$F(z, u) = \frac{Q(u, z)}{N(u, z)} = \frac{\prod_{i=1}^{c}(u - u_i)}{-z s_d \prod_{j=1}^{c+d}(u - u_j)}$$

En $u = 0$ (excursions) et en $u = 1$ (marches), on obtient ainsi une série génératrice algébrique (prouve une généralisation de la conjecture de Pinzani). $\square$

# Marches sur $\mathbb{N}$ à pas avant bornés et arrière "continus"

$[(k) \to (1)(2)(3)\dots(k-4)(k-1)(k+5)]$

$$L : u^k \mapsto u^0 + u^1 + \cdots + u^{k-1}$$

$$L := \frac{U}{1-u} + \frac{1}{1-u}$$

$$F(z,u)(1 + \frac{z}{1-u} - zP(u)) = 1 + \frac{z}{1-u}F(z,1)$$
$$-z\sum_{j=0}^{c-1} c_j(u)\partial_u^j F(z,0).$$

$$N(u,z) := -u^c(1-u) - u^c z + u^c(1-u)P(u).$$

- une branche équivalente à 1 en $z = 0$, que nous noterons $u_0$ ;
- $c$ petites branches équivalentes à $z^{1/c}$ en $z = 0$, que nous noterons $u_1, \dots, u_c$ ;
- $d$ branches infinies en 0, que nous noterons $v_1, \dots, v_d$.

**Théorème 5.** *Une marche sur $\mathbb{N}$ à pas arrière "continus" et à pas avant bornés a une série génératrice des excursions algébrique et elle vaut*

$$\frac{(-1)^{c+1}}{s_d z} \prod_{i=0}^{c} u_i.$$

**Théorème 6.** *Une marche sur $\mathbb{N}$ à pas arrière continus et à pas avant bornés a une série génératrice des marches algébrique et elle vaut*

$$\frac{-1}{z} \prod_{i=0}^{c} (1 - u_i).$$

**Exemples :**

Catalan $[(2), (k) \rightarrow (2)(3) \ldots (k)(k+1)]$

arbres $t$-aires $[(t), (k) \rightarrow (t) \ldots (k) \ldots (k+t-1)]$

Motzkin $[(1), (k) \rightarrow (1)(2) \ldots (k-1)(k+1)]$

$$F(z,1) = \frac{1 - z - \sqrt{1 - 2z - 3z^2}}{2z^2} = \{1, 1, 2, 4, 9, 21\}$$

Schröder $[(3), (k) \rightarrow (3) \ldots (k)(k+1)^2]$

$$F(z,1) = \frac{1 - 3z - \sqrt{1 - 6z + z^2}}{4z^2} = \{1, 3, 11, 45, 197\}$$

# Marches en dimension 2: des chemins de Dyck bien cavaliers

Les pas du cavalier sont codés par le polynôme caractéristique

$$P(u,v) := uv^2 + uv^{-2} + u^2v + u^2v^{-1}$$

On a la récurrence

$$f_{n+1} := Pf_n - \{u^{<0}v^{<0}\}Pf_n$$

d'où l'équation fonctionnelle

$$F(z,u,v) = 1 + z\big(P(u,v)F(z,u,v)$$

$$-uv^{-1}\partial_v F(z,u,0) - (u^2v^{-1} + uv^{-2})F(z,u,0)\big)$$

En annulant le noyau, on obtient

$$\begin{cases} v_0(u,z) = +\sqrt{z}\sqrt{u} + \cdots \\ v_1(u,z) = -\sqrt{z}\sqrt{u} + \cdots \end{cases}$$

et finalement

$$F(z,u,v) = \frac{v^2 + v_0 v_1 - v v_0 - v v_1}{v^2 - z(u^2 v^3 + u v^4 + u + u^2 v)}$$

"nombre de chemins allant de $(0,0)$ à $(n,0)$" :

$$a_n = [u^n v^0] F(1,u,v).$$

$$F(1,u,0) = \frac{1 + 2u + \Delta}{4u} - \frac{\sqrt{2 + 2\Delta - 8u^2 + 4u\Delta}}{4u}$$

$$\Delta := \sqrt{1 - 4u + 4u^2 - 4u^4}$$

$$a_n \sim (1 + \sqrt{3})^n \approx 2.73^n \quad \square$$

"séquences de $n$ mouvements de cavalier
partant de $(0,0)$ et qui reviennent sur $y = 0$"

$$b_n = [z^n v^0] F(z,1,v)$$

$$F(z,1,0) = \frac{1 + 2z + \Delta}{4z} - \frac{\sqrt{2 + 2\Delta - 12z^2 + 4z\Delta}}{4z}$$

$$\Delta := \sqrt{1 - 4z}$$

$$b_n \sim 4^n \quad \square$$

## Limitation par une droite de pente rationnelle

Définition (langages de Dyck généralisés) : On prend un alphabet fini, on donne à chaque lettre une valeur entière positive ou négative et on forme un langage avec des conditions de positivités similaires à celles des mots de Dyck (valeur totale nulle, valeur positive pour tous les facteurs gauches).

**Théorème 7.** *La série génératrice des chemins de Dyck généralisés est algébrique.*

Preuve.

Pour deux lettres $\Leftrightarrow$ marches comprises entre $y = 0$ et $y = \frac{p}{q}x$. Équivalent à une marche en dimension 1 de polynôme caractéristique

$$P(u) := P(u^{-p}, u^q)$$

Cas général : polynôme caractéristique du poids des lettres et théorème "pas bornés". $\square$

# Opérateur

Tout marche bien quand on trouve un opérateur linéaire.

Dérivation (intégration). $\partial : u^k \mapsto ku^{k-1}$
Substitution (Mahler). $\sigma_\alpha : u \mapsto u^\alpha$
Évaluation en 1. $U = \sigma_0 : u \mapsto 1$
Évaluation en 0. $Z = \sigma_\infty : u \mapsto 0$
Shift. $S : u^k \mapsto u^{k+1}$
Substitution (shift additif). $u \mapsto u + a$
Mise à zéro sur $E$. $u^k \mapsto 0$ si $k \in E$
Extraction. $[u^k]$, $\{u^k\}$

**Applications**

- Sauts "linéaires" avec "multiplicité linéaire" $(k) \rightarrow (ak+b)^{(ck+d)}$ se code par $(ck+d)u^{ak+b}$ et donc $L = cS^b\sigma_a S\partial + dS^b\sigma_a$. $\square$

- Pour une s.g. donnant l'aire, on veut :

$$L(u^i v^j q^k) = u^{i+a} v^{j+b} q^{k+aj+\frac{ab}{2}}.$$

C'est okay avec l'opérateur de substitution $\sigma : v \mapsto vq^a$ et en prenant

$$L := S_q^{\frac{1}{2}} S_u S_v \sigma + S_q^{-\frac{1}{2}} S_u S_v^{-1} \sigma \quad \square$$

# Nombre moyen d'effleurements

$$f_{n+1} = P(u) f_n - \{u^{<0}\} P(u) f_n$$

$$-[u^0] P(u) f_n + t[u^0] P(u) f_n$$

Exemple avec les mots de Dyck $(P(u) = u + \frac{1}{u})$ :

$$F(z,u,t) = \frac{u^2(1-t)(u_1 - z) + z u_1}{((1-t)u_1 + zt)(zu^2 - u + z)}$$

Le nombre moyen d'effleurements est

$$\mu_{2k} = \frac{[z^{2k}]\partial_t F(z,0,1)}{[z^{2k}]F(z,0,1)} = \frac{[z^{2k}]\frac{u_1(z)(u_1(z)-z)}{z^2}}{[z^{2k}]\frac{u_1(z)}{z}}$$

$$= \frac{\frac{\binom{2(k+1)}{k+1}}{k+2} - \frac{\binom{2k}{k}}{k+1}}{\frac{\binom{2k}{k}}{k+1}} = \frac{(k+1)(2k+1)(2k+2)}{(k+2)(k+1)^2} - 1$$

$$= 3\frac{k}{k+2} = 3\sum_{p\geq 0} \frac{(-1)^p 2^p}{k^p} = 3 - \frac{6}{k} + \frac{24}{k^2} + O(\frac{1}{k^3}).$$

$\square$

# Asymptotique des excursions

Combien y a-t-il, pour $n$ “grand”, d’excursions de longueur $n$ ?

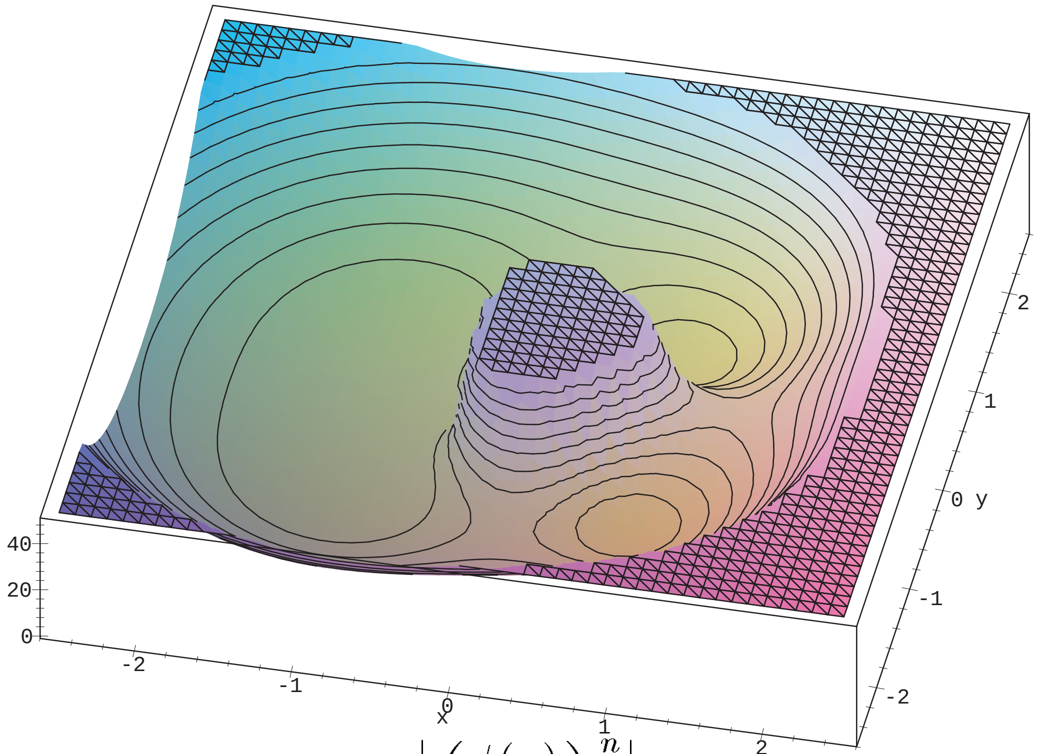
**Théorème 8.** *Asymptotique en $\frac{A}{\rho^n\sqrt{\pi n^3}}$. $A$ et $\rho$ (le rayon de convergence) sont algébriques.*

Preuve.
$F = \prod U_i$. Or singularité(s) de $U_i$ dicte(nt) la croissance des coeff de $F$.
$c = \frac{u\phi'(u)}{\phi(u)}$ une unique racine positive $\tau$.
De plus $u(\rho) = \tau$. Par ailleurs, lemme :

$$\forall x > 0 \ : \phi(x)(x\phi''(x) + \phi'(x)) > x(\phi'(x))^2$$

$$\text{d'où} \quad u - \tau = d_1\sqrt{\rho - z} + d_2(\rho - z) + \dots \qquad \square$$

Autre preuve. Méthode du col.
Cauchy nous dit $f_n = \int_\gamma \frac{F(z)}{z^{n+1}}dz$
Idée : jouer avec le contour $\gamma$ et obtenir l’asymptotique quand $n$ est grand. $\square$




$$\left|\left(\frac{\phi(u)}{u^c}\right)^n\right|$$

(Ci-dessus $\phi = 1 + 2u + u^4 + u^5$ et $c = 3$, *i.e.* $(k) \rightarrow (k-3)(k-2)^2(k+1)(k+2)$).

Quand $n$ est grand, tout est concentré autour du point col.

# Fractions continuées

naturel de par la traduction structure combinatoire /analyse :
**F=Seq(A+Seq(B+Seq(...))**

$$\Leftrightarrow F(z) = \cfrac{1}{1 - A(z) - \cfrac{1}{1 - B(z) - \ddots}}.$$

Nos s.g. étaient des s.g. ordinaires, mais pas de problème si on recourt aux s.g. exponentielles :

$$\tilde{F}(z,u) = \sum f_n(u)\frac{z^n}{n!} = \sum \tilde{f}_n(u) z^n$$

$$\partial_z \tilde{F}(z,u) = L\tilde{F}(z,u).$$

On a ainsi de jolies formes closes pour :
Arrangements $[2,(k) \rightarrow (k)(k+1)^{k-1}]$

$$\tilde{F}(z,u) = \frac{\exp(z)}{1-uz}$$

Laguerre $[2,(k) \rightarrow (k+1)^{k-1}(k+2)]$

$$\tilde{F}(z,u) = \frac{\exp(\frac{u^2 z}{1-uz})}{1-uz}$$

Hermite $[2, (k) \to (k-1)^{k-1}(k+1)]$
(involutions en marquant les points fixes)

$$\tilde{F}(z,u) = \exp(zu + \frac{z^2}{2})$$

$$F(z,0) = \cfrac{1}{1 - \cfrac{z^2}{1 - \cfrac{2z^2}{1 - \cfrac{3z^2}{1 - \cfrac{4z^2}{\ddots}}}}}$$

Premiers termes :$\{1, 0, 1, 0, 3, 0, 15, 0, 105\}$

Les suivantes ne sont plus seulement transcendantes mais même plus holonomes :
Stirling $[2, (k) \to (k)^{k-1}(k+1)]$
(nb de partages de $n$ en $k$ non vide)

$$\tilde{F}(z,u) = \exp(u(\exp(z) - 1))$$

$$F(z,1) = \cfrac{1}{1 - z\cfrac{z^2}{1 - 2z\cfrac{2z^2}{1 - 3z - \cfrac{3z^2}{1 - 4z - \cfrac{4z^2}{\ddots}}}}}$$

Premiers termes :$\{1, 2, 5, 15, 52, 203\}$
Bessel $[0, (k) \to (k-1)(k)^k(k+1)]$
(permutations sans croisements)

$$F(z,0) = \frac{1}{1 - z - z^2 B(z)}$$

$$F(z,0) = \cfrac{1}{1 - z - \cfrac{z^2}{1 - z - \cfrac{z^2}{1 - 2z - \cfrac{z^2}{1 - 3z - \cfrac{z^2}{\ddots}}}}}$$

Premiers termes :$\{1, 1, 2, 4, 9\}$